Fermilab U.S. DEPARTMENT OF ENERGY | Office of Science

# Programming Language Experience Survey

Marc Paterno

*5 Dec, 2019*

# Overview

- The goal of the survey was to identify groups of SCD people with expertise in C++ or Python languages.
  - find those with expertise that is not widely known
  - not a means to pigeonhole anyone
- A short survey does not paint a complete picture of anyone's skils
  - does not identify those who *want* to develop greater expertise
  - does not measure many types of expertise (*e.g.* system design skill)

# The surveys

- Questions were designed to probe a variety of language features and techniques.
- *Yes* answers always correspond to experience with more features.
- Don't start with assumption of the relative importance of each question.

# Analysis

- We have 129 responses to the C++ survey, 127 to the Python survey.
- I used *K-means clustering*, using both 2 and 3 clusters.
- For each number of clusters, we do this with and without using the response for self-identified experts.
- Algorithm doesn't identify which cluster correponds to highest (*expert*) level of experience; that is added by me.
- I am happy to share all the details of the analysis with anyone interested.

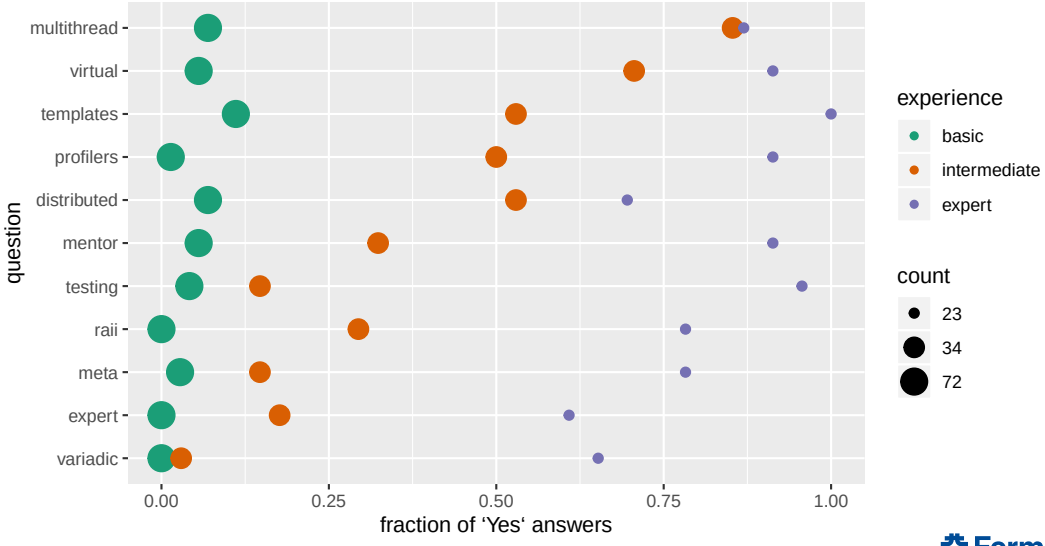
# C++ results

- There are 1 classifications that disagree between the `c2` and `c2ne` clustering.
- 2 between the `c3` and `c3ne` classifying.

| c2 | c2ne | c3 | c3ne |
|---|---|---|---|
| expert | basic | intermediate | intermediate |
| expert | expert | intermediate | expert |
| expert | expert | intermediate | expert |

- The `c3ne` results appear best.
- Under this assumption we have 23 experts.

# C++ question quality assessment






Fermilab

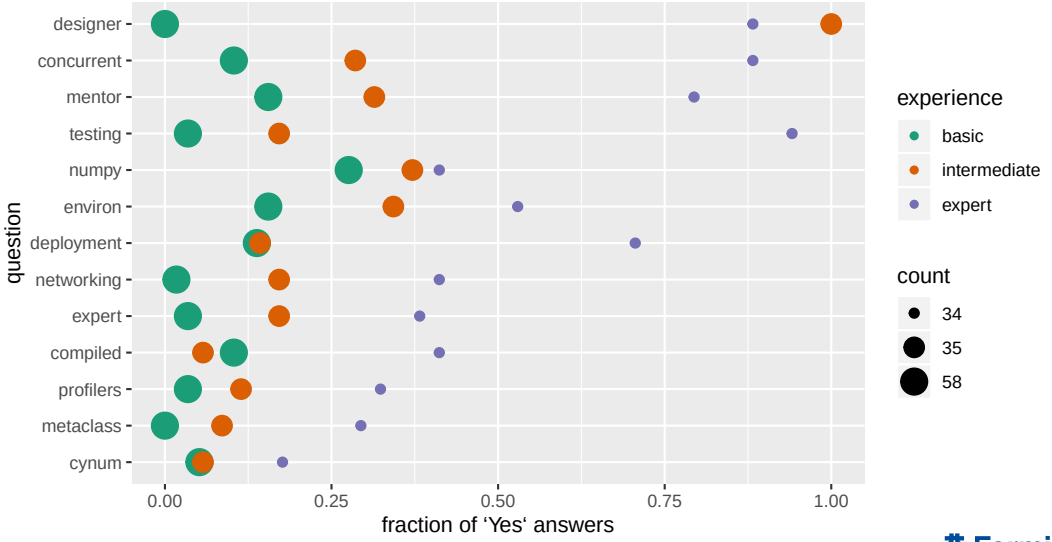
# Python results

- Initial inspection of question results indicates some were identifying users of Python for data science; I ignored them in clustering.
- 9 disagreements between `c2` and `c2ne`
- 4 between `c3` and `c3ne`
- I went with `c3ne` results again.
- Under this assumption we have 34 experts.

# Python disagreement details

| c2 | c2ne | c3 | c3ne |
|---|---|---|---|
| expert | basic | intermediate | intermediate |
| expert | expert | intermediate | expert |
| expert | basic | intermediate | intermediate |
| expert | basic | intermediate | intermediate |
| expert | basic | expert | basic |
| expert | basic | intermediate | intermediate |
| expert | basic | intermediate | intermediate |
| expert | expert | expert | intermediate |
| expert | basic | intermediate | intermediate |
| expert | expert | basic | expert |
| expert | basic | intermediate | intermediate |
| expert | basic | intermediate | intermediate |

# Python question quality assessment






Fermilab

# What overlap is their between the groups of experts?

| expertise | count |
|---|---|
| both | 8 |
| c++ | 15 |
| py | 26 |

## What is next? The *start* of a discussion

- Previous C++ training (*e.g.* Downing's course) has been aimed at basic level skill.
  - I think Glenn's course was good, and if we have a budget to do so, we should repeat it.
  - We can work with him to tweak it to be even better.
  - We should encourage people *in the SCD*, not just experimenters, to participate.
- Is there a budget to offer training to help move the *intermediate* level toward *expert*?
  - I have started to look at some professional training courses.
  - Sent messages to a few people at CERN, where some courses have been given.
  - Will also send questions to IRIS-HEP about possibilities.
- I would like to also try something less formal:
  - biweekly topical "seminar" / "discussion group"
  - Make use of freely-available online materials
  - View *and discuss*, in perhaps a 90-minute session
  - Invite any interested parties.
  - What would be needed to get project leaders and department heads to encourage participation?
- **What ideas do you have?**